Journal of Mathematical Analysis and Applications **252,** 696–709 (2000)
doi:10.1006/jmaa.2000.7095, available online at http://www.idealibrary.com on IDEAL®

Invariant Measures of Set-Valued Maps

Zvi Artstein[1]

Department of Mathematics, The Weizmann Institute of Science

Submitted by William F. Ames

Received February 26, 1999

1. INTRODUCTION AND RESULTS

Invariant measures of maps from a space into itself play a crucial role in mathematical analysis and its applications. The standard definition of an invariant measure requires that the function be univalued. Recent applications (see e.g., [2]) have raised the need for an analogous notion of invariant measures for multifunctions, namely, maps that assign to each point a subset of the space. We display in this note five natural suggestions for the needed notion and examine when they are equivalent. In fact, we find that four of our definitions are equivalent when the underlying space is a separable and complete metric space, and the fifth one is equivalent to them if, in addition, the space is locally compact and the set-valued map has a closed graph.

Let X be a complete separable metric space with its Borel structure. Recall that a probability measure p on X is invariant with respect to the measurable function f: $X \to X$ if

$$p(B) = p\big(f^{-1}(B)\big) \tag{1.1}$$

for every Borel set $B \subseteq X$ (here, as usual, $f^{-1}(B) = \{x : f(x) \in B\}$).

The set-valued functions we consider are maps F which assign to each point $x \in X$ a closed subset $F(x) \subseteq X$. We assume that F is measurable. The measurability is defined for set-valued maps, say G, from a space, say Y, into the closed subsets of another space, say Z, by the statement: for

[1] Incumbent of the Hettie H. Heineman Professorial Chair in Mathematics. Research was supported by a grant from the United States–Israel Binational Science Foundation (BSF), by the MINERVA Foundation (Germany), and by INTAS (Belgium).

0022-247X/00 $35.00

every closed set $C \subseteq Z$, the set $G^{-}(C) = \{y : G(y) \cap C \neq \emptyset\}$ is measurable. See Castaing and Valadier [6] or Rockafellar [11] for background on measurable multifunctions. Note that in this paper we are dealing with the case where Y and Z coincide.

We now display the promised five properties that probability measures on X may have in relation to the set-valued mapping F. Each of these properties offers a natural definition of an invariant measure of the set-valued map. When needed, the statement of a property is preceded by some notions and terminology. Following the list we state the equivalence theorems, comment on the sources of our list, and give some references.

Property I. A probability measure p on X is a subinvariant measure of F if

$$p(C) \leq p(F^{-}C) \tag{1.2}$$

for every closed set $C \subseteq X$.

Another natural definition would be to have p an invariant measure with respect to a selection of F. Example 4.1 shows the limitations of this definition. Instead we offer the following two properties. A Markov transition function μ from X to itself is a map which assigns to each x in X a probability measure $\mu(x)$ on X and is such that μ is measurable with respect to the metric of weak convergence. Let p be a probability measure on X. The probability measure induced from p by the Markov transition function μ, denoted $\mu(p)$, is given by

$$\mu(p)(B) = \int \mu(x)(B) p(dx). \tag{1.3}$$

The support of a probability measure q is denoted by $\operatorname{supp} q$.

Property II. A probability measure p is a Markov invariant measure of the set-valued F if there exists a Markov transition function μ on X such that $\operatorname{supp} \mu(x) \subseteq F(x)$ for p-almost every x and such that $\mu(p) = p$.

We consider now the space of closed subsets of X. It can be endowed with a topological structure where the sets

$$I_C = \{D : D \subseteq X \text{ is closed and } D \cap C \neq \emptyset\}, \tag{1.4}$$

for C closed in X, are closed. For a general theory see Beer [4]. Here we need only to know when two random maps into this space have the same distribution. Namely, let p be a probability measure on X and F be a measurable set-valued map. Let H be a set-valued map from a metric space Y into the closed subsets of X, and suppose that Y is endowed with its Borel field, q is a probability measure on Y, and H is measurable. We say that H has the same distribution as F, or that H is a version of F, if

$$p\big(F^{-1}(I_C)\big) = q\big(H^{-1}(I_C)\big)$$

for every closed subset C of X. As usual, if h is a measurable function from Y into X, and q is a probability distribution on Y, then h induces from q a probability measure $h(q)$ on X, given by $h(q)(B) = q(h^{-1})(B)$; compare this with (1.3).

Property III. The probability measure p on X is selectable with respect to the set-valued mapping F if there exists a version H of F, defined on a metric space Y endowed with a probability measure q, and a measurable selection h of H, such that $p = h(q)$.

Consider the collection X^∞ of double infinite sequences

$$\chi = (\ldots, x_{-2}, x_{-1}, x_0, x_1, x_2, \ldots) \tag{1.5}$$

with coordinate-wise convergence; a metric can be constructed on X^∞ which makes it a complete and separable metric space. Let P be a probability distribution on X^∞. The projection of P on X is the probability measure p given by

$$p(B) = P(\{\chi : x_0 \in B\}). \tag{1.6}$$

On X^∞ consider the translation map T, namely

$$T(\chi)_i = x_{(i+1)} \qquad \text{for all } i. \tag{1.7}$$

Then T is a one-to-one continuous map. In particular, the standard notion of an invariant measure (see (1.1)) applies to the mapping T on X^∞. Denote by Ξ the collection of points $\chi \in X^\infty$ with the property

$$x_i \in F(x_{(i-1)}) \qquad \text{for all } i. \tag{1.8}$$

It follows from [6, Theorem III.30] that the set $\{(x, y) : y \in F(x)\}$ is a Borel set in $X \times X$, hence it is easy to see that Ξ is a Borel subset of X^∞.

Property IV. The probability measure p on X is a projectional invariant measure of the set-valued map F if it is the projection of a probability measure P on X^∞ which satisfies $P(\Xi) = 1$ and is such that P is invariant with respect to the translation map T.

Consider a sequence $(x_1, x_2, x_3, \ldots)$ such that $x_{(i+1)} \in F(x_i)$ for all $i = 1, 2, \ldots$. Such a sequence will be called a trajectory of the set-valued map F. The occupational measure of a finite segment $(x_1, \ldots, x_N)$ of a trajectory will be denoted by p_N (when we use this notion the trajectory will be known, so we suppress the segment itself from the notation) and it is given by

$$p_N(B) = \frac{1}{N} \sharp \{i : 1 \leq i \leq N \text{ and } x_i \in B\}, \tag{1.9}$$

where $\sharp$ indicates the cardinality of a set. Recall the weak convergence of probability measures (see, e.g., Billingsley [5]). Recall the convex structure on the space of probability measures generated by the operation $(\alpha p + (1-\alpha)q)(B) = \alpha p(B) + (1-\alpha)q(B)$ for $0 \leq \alpha \leq 1$.

Property V. The probability measure p on X is an occupational invariant measure of the set-valued map F if it is in the closed convex hull of the individual occupational invariant measures of F (where a probability measure p is an individual occupational probability measure of the set-valued function F if there exists a trajectory of F whose occupational measures p_N converge to p in the weak convergence of measures).

THEOREM A. *Properties I–IV are equivalent when X is a complete separable metric space and F is a measurable set-valued map.*

THEOREM B. *Property V is equivalent to* (*each of*) *Properties I–IV if X is a locally compact complete separable metric space and F has a closed graph.*

Sources. The notion displayed in Property I was offered in Aubin *et al.* [3] as the definition of an invariant measure. The Poincaré recurrence theorem was established in [3] for such invariant measures under the assumption that X is compact and F is continuous. Properties IV and V were proposed by the author in [2] (not being aware of [3]), however, for set-valued dynamics generated by differential inclusions, where their equivalence was established. A construction similar to the lifting used in Property IV was employed by Colonius and Kliemann [7] in a control systems framework. In [2] the two properties were applied to problems of singular perturbations. Property II was proposed to the author by Vershik during a visit to the Weizmann Institute. It has its roots in [13]. Vershik conjectured the equivalence of Properties II and IV. The equivalence of Properties II and III in the case in which X is locally compact follows from the paper [1] of the author; the more general related results by Ross [12] and by Živaljević [14] may help in establishing our results under more general conditions.

In view of the stated results we could choose each of the first four properties as a general underlying definition of an invariant measure of a set-valued function and have the other properties serve as characterizations. We prefer to define that a probability measure is an invariant measure of the set-valued map F if it satisfies one of the first four properties (or Property V if the assumptions of Theorem B are met).

The proof of Theorem A is presented in the next section. It follows a series of propositions establishing, respectively, the following implications: I $\Rightarrow$ II, II $\Rightarrow$ III, III $\Rightarrow$ I, IV $\Rightarrow$ I, and II $\Rightarrow$ IV. In Section 3 we present two propositions establishing, under the additional conditions of Theorem B, that V $\Rightarrow$ I and IV $\Rightarrow$ V; these verify Theorem B. In the closing section we display some comments, counterexamples, and applications.

2. PROOF OF THEOREM A

PROPOSITION 2.1. *A subinvariant measure is Markov invariant.*

Proof. Our first step is to construct an approximation to the problem where the underlying space is compact and the set-valued map is continuous. The proof of the original problem will be obtained through a limit argument.

Let p be the subinvariant measure. Being a probability measure on a complete separable space, p is tight (see Billingsley [5]); namely, for every $\varepsilon > 0$, a compact set $K_\varepsilon \subseteq X$ exists such that $p(K_\varepsilon) > 1 - \varepsilon$.

Consider the set-valued function F_ε from K_ε into the closed subsets of K_ε given by $F_\varepsilon(x) = F(x) \cap K_\varepsilon$. The multifunction F_ε is measurable, as it follows, e.g., from Rockafellar [11, Theorem 1M]. Since K_ε is compact, the Lusin theorem (see e.g. [11, Theorem 1F]) implies that there exists a compact subset L_ε of K_ε such that $p(L_\varepsilon) > 1 - 2\varepsilon$ and such that on L_ε the mapping F_ε is continuous with respect to the Hausdorff metric on compact sets (see e.g. [6] for the latter notion).

The p-measure of K_ε and L_ε is less than 1. We consider therefore the spaces $\hat{K}_\varepsilon = K_\varepsilon \cup \{k_\varepsilon\}$ and $\hat{L}_\varepsilon = L_\varepsilon \cup \{l_\varepsilon\}$ where k_ε and l_ε are two auxiliary points. We consider the space $\hat{K}_\varepsilon$ as a measure space where the measure on K_ε is determined by p and the measure of $\{k_\varepsilon\}$ is $1 - p(K_\varepsilon)$. We denote the resulting probability measure on $\hat{K}_\varepsilon$ by $p_{\varepsilon,k}$. Likewise, we define the probability measure $p_{\varepsilon,l}$ on $\hat{L}_\varepsilon$ by letting $p_{\varepsilon,l}$ be equal to p on L_ε and by $p_{\varepsilon,l}(\{l_\varepsilon\}) = 1 - p(L_\varepsilon)$.

At this point we have two compact metric spaces $\hat{K}_\varepsilon$ and $\hat{L}_\varepsilon$ and probability distributions defined on each of them. Consider the relation $R \subseteq \hat{L}_\varepsilon \times \hat{K}_\varepsilon$ given by $(y, x) \in R$ if either $(y, x) \in L_\varepsilon \times K_\varepsilon$ and $x \in F(y)$ or whenever $y = l_\varepsilon$ or $x = k_\varepsilon$ (we suppress the dependence of the relation R on ε). Since F is a continuous multifunction on $\hat{L}_\varepsilon$ it follows that R is a closed subset of $\hat{L}_\varepsilon \times \hat{K}_\varepsilon$.

We claim that since p is subinvariant with respect to the multifunction F, it follows that for every measurable C in $\hat{K}_\varepsilon$ the following inequality holds:

$$p_{\varepsilon,k}(C) \le p_{\varepsilon,l}(\{y : (y,x) \in R \text{ and } x \in C\}). \tag{2.1}$$

The claim follows directly from (1.2) and the fact that $\{l_\varepsilon\} \times \hat{K}_\varepsilon$ is a subset of R. By [1, Theorem 3.1] it follows that a Markov transition map exists from $\hat{L}_\varepsilon$ to $\hat{K}_\varepsilon$, denoted $\mu_\varepsilon(\cdot)$, such that $\operatorname{supp} \mu_\varepsilon(y) \subseteq F(y)$ for every $y \in \hat{L}_\varepsilon$ and such that $\mu_\varepsilon(p_{\varepsilon,k}) = p_{\varepsilon,l}$, namely, $p_{\varepsilon,k}$ is induced by μ_ε from $p_{\varepsilon,l}$ (see (1.3) for the notation).

The preceding paragraphs establish the promised approximation (note that the restrictions of $p_{\varepsilon,l}$ and $p_{\varepsilon,k}$ to L_ε and K_ε, respectively, coincide with p). At this point we start with the limit arguments. To this end we identify the measure-valued map $\mu_\varepsilon(\cdot)$ with the probability measure $\boldsymbol{\mu}_\varepsilon$ on $L_\varepsilon \times K_\varepsilon$ obtained by direct integration. Namely, on rectangles, say on $E \times C$, the measure is given by

$$\boldsymbol{\mu}_\varepsilon(E \times C) = \int_E \mu_\varepsilon(y)(C)\, p_{\varepsilon,l}(dy). \tag{2.2}$$

Consider now a sequence $\varepsilon_i \to 0$. We claim that the sequence of measures $\boldsymbol{\mu}_{\varepsilon_i}$ ((defined on $(X \cup N) \times (X \cup N)$ where $N = \{1,2,3,\dots\}$) is tight. Indeed, given an $\varepsilon > 0$, an easy calculation shows that $\boldsymbol{\mu}_{\varepsilon_i}(L_\varepsilon \times K_\varepsilon) \ge 1 - 2(\varepsilon + \varepsilon_i)$. This implies the tightness.

The tightness implies that a subsequence exists, say $\boldsymbol{\mu}_{\varepsilon_j}$, which converges weakly, say, to $\boldsymbol{\mu}$. Since the respective weights that the measures $p_{\varepsilon,k}$ and $p_{\varepsilon,l}$ assign to N tend to zero as ε tends to zero, it follows that the measure $\boldsymbol{\mu}$ is supported on $X \times X$.

Since for every j the marginals of $\boldsymbol{\mu}_{\varepsilon_j}$ on L_{ε_j} and on K_{ε_j} are both equal to p, and since both $p(L_\varepsilon)$ and $p(K_\varepsilon)$ converge to 1 as ε tends to zero, it follows that the marginals of $\boldsymbol{\mu}$ on both coordinates are equal to p. In particular, $\boldsymbol{\mu}$ can be disintegrated with respect to the y-coordinate endowed with p, resulting in a measure-valued function, say $\mu(\cdot)$. We claim that the measure-valued map $\mu(\cdot)$ is the desired Markov transition map, namely, $\mu(y)$ is a probability measure satisfying $\operatorname{supp} \mu(y) \subseteq F(y)$, for p-almost every y, and $\mu(p) = p$. The latter equality is satisfied since $\mu(\cdot)$ is obtained by disintegrating a measure whose two marginals coincide with p. To verify that $\mu(\cdot)$ is a probability measure note that since for each i the measure $\boldsymbol{\mu}_{\varepsilon_i}$ is obtained by (2.2) and since $\mu_{\varepsilon_i}(y)$ is known to be a probability measure, it follows that $\mu(y)$ is indeed p-almost everywhere a probability measure. It remains to verify that $\operatorname{supp} \mu(y) \subseteq F(y)$ for p-almost every y. To this end, assume without loss of generality that $\Sigma\varepsilon_j$ is finite. For j_0 large enough, the intersection of all the sets L_{ε_j}, for $j \geq j_0$, has p-measure close to 1, say, greater than or equal to $1 - \delta$. Denote this intersection by L_δ. Likewise, the intersection K_δ of all K_{ε_j} for $j \geq j_0$ also has p-measure greater than $1 - \delta$. The set $L_\delta \times K_\delta$ is compact, and the restriction to $L_{\varepsilon_j} \times K_{\varepsilon_j}$ of each of the $\boldsymbol{\mu}_{\varepsilon_j}$ is supported on R. Hence the restriction of $\boldsymbol{\mu}$ to $L_\delta \times K_\delta$ is also supported on R. Finally, the restriction of the graph of F to the same set is compact and equals the restriction of R to the same set. This implies that the restriction of supp $\mu(y)$ to K_δ for $y \in L_\delta$ is in $F(y)$ for p almost every y. Since δ is arbitrarily small, and since the μ-measure of $L_\delta \times K_\delta$ tends to 1 as δ tends to 0, the proof is complete.

PROPOSITION 2.2. *A Markov invariant measure is selectable.*

Proof. The result follows standard construction arguments in probability theory (a special case was established in [1, p. 318]). For completeness we sketch the proof.

Let p be a probability measure on X which is Markov invariant with respect to the set-valued mapping F, and let $\mu(\cdot)$ be the Markov transition function which leaves p invariant. The auxiliary space on which we construct the version of F is $X \times [0,1]$, where $[0,1]$ is endowed with the Lebesgue measure λ and the version itself, say G, is given by $G(x,t) = F(x)$. The goal is to define a function g: $X \times [0,1] \to X$ which is measurable and such that, for every x_0 fixed, the measure induced from $g(x_0, \cdot)$ by λ is equal to $\mu(x_0)$. Since supp $\mu(x) \subseteq F(x)$ for p-almost every x, it is clear that $g(x,t) \in G(x,t)$ for $(p \times \lambda)$-almost every (x,t), hence the pair (G, g) would indeed verify that p is selectable.

For the construction itself, it does not matter that X is both the domain and the range of $\mu(\cdot)$ and that $\mu(p) = p$; we therefore denote the range by Z and denote by q the measure induced from p by $\mu(\cdot)$.

Since Z is separable and complete, it follows that q is tight, and in particular that there exists a sequence $A_{1,i}$ for $i = 1, 2, \ldots,$ of disjoint measurable subsets of Z, each with a diameter less than or equal to, say, 1; a compact closure; and such that $q(\cup A_{1,i}) = 1$. Successively on the index k, we define a sequence $A_{k,i}$ for $i = 1, 2, \ldots$ of disjoint measurable subsets of Z, each with diameter less than or equal to $1/k$ and with the additional property that each $A_{k-1,j}$ is the union of a finite string $A_{k,i_1} \cup A_{k,i_1+1} \cup \cdots \cup A_{k,i_1+j_1}$ for some i_1 and j_1. Such a partition is possible since each $A_{1,i}$ has a compact closure.

For each k define now $\theta_k(x, i) = \mu(x)(A_{k,1} \cup \cdots \cup A_{k,i})$, for $i = 1, 2, \ldots$. Since μ is measurable it is clear that θ_k is measurable. For each i let $z_{k,i}$ be a point in $A_{k,i}$. Define $g_k(x,t) = z_{k,i}$ if $\theta_k(x, i-1) < t \leq \theta_k(x,i)$, when $\theta_k(x, 0) = 0$. It is clear that g_k is measurable and that, for every x fixed, the distribution induced from λ by $g_k(x, \cdot)$ is close in the weak convergence of measures to $\mu(x)$. The nested structure of $A_{k,i}$ implies that for each (x, t) the sequence $g_k(x, t)$ is a Cauchy sequence in Z. Hence the sequence g_k converges pointwise (in fact, uniformly, due to the estimate $1/k$). The limit, say g, clearly satisfies the desired properties. This completes the proof.

PROPOSITION 2.3. *A selectable measure is subinvariant.*

Proof. Let p be a selectable probability measure of the set-valued map F and let (Y, q, H, h) be the data verifying that, namely, Y is a complete separable metric space with q a probability measure on it; H is a measurable set-valued map from Y into closed subsets of X which is a version of F; and h: $Y \to X$ is a measurable selection of H such that $h(q) = p$. We claim that for every closed set C of X,

$$\begin{aligned} p(C) &= h(q)(C) = q(\{y: h(y) \in C\}) \le q(\{y: H(y) \cap C \neq \varnothing\}) \\ &= p(F^- C). \end{aligned} \tag{2.3}$$

The first equality follows from the selectability. The second equality is the definition of $h(q)(C)$. The inequality is justified since $h(y) \in C$ implies $H(y) \cap C \neq \varnothing$. The last equality follows since H is a version of F. In particular, (2.3) verifies that (1.2) holds; this completes the proof.

A string similar to (2.3) would as easily prove that a Markov invariant measure is subinvariant.

PROPOSITION 2.4. *A projectional invariant measure is subinvariant.*

Proof. Let p be a projectional invariant measure on X and let P be the invariant measure on X^∞ whose support is on Ξ and whose projection is p. Then for a closed subset C of X the following holds (we maintain the notation used in (1.5)):

$$\begin{aligned} p(C) &= P(\{\chi : x_0 \in C\}) \le P(\{\chi : x_{-1} \in F^- C\}) \\ &= P(\{\chi : x_0 \in F^- C\}) = p(F^- C). \end{aligned} \tag{2.4}$$

The first and the last equalities follow from the definition of the projection. The inequality follows since P is supported on Ξ, hence for P-almost every χ if $x_0 \in C$ then $x_{-1} \in F^-(C)$. The remaining equality follows since P is invariant under the translation operator T. Since (2.4) verifies (1.2), the proof is complete.

PROPOSITION 2.5. *A Markov invariant measure is projectional invariant.*

Proof. Let p be a Markov invariant probability measure of the set-valued mapping F, and let $\mu(\cdot)$ be the Markov transition function which verifies that. The transition function maps X to probability measures on X. The probability measure p and the transition function $\mu(\cdot)$ induce in a standard way a probability measure on any finite product

$$X \times X \times \cdots \times X. \tag{2.5}$$

The construction can be done inductively. On X the measure is $p = P_1$. Suppose that on an n-string $Y = X \times X \times \cdots \times X$ the measure is well-defined, say P_n. Define the transition measure from Y into probability measures on X by

$$\mu_n(x_1, \dots, x_n) = \mu(x_n). \tag{2.6}$$

The measure on the $(n + 1)$ string $Y \times X$ is then determined by its values on rectangles $A \times B$ given by

$$P_{n+1}(A \times B) = \int_A \mu_n(y)(B) P_n(dy). \tag{2.7}$$

The values of P on rectangles determine the probability measure on the product (see Neveu [9, Proposition II.2.1]; we have already used this construction in (2.2)).

The preceding construction defines a probability measure on each finite substring of X^∞. Since p is Markov invariant with respect to μ, it is clear that the measure on a string is compatible with the measure on any substring; namely, if A_i are subsets of X then $P_n(A_1 \times \cdots \times A_{n-1} \times X) = P_{n-1}(A_1 \times \cdots \times A_{n-1})$. The compatibility implies that the measures on the strings induce (uniquely) a probability measure, say P_∞, on X^∞. See, e.g., [9, Proposition III.3.3].

We claim that the probability measure P_∞ on X^∞ is the desired probability measure; namely, it is invariant with respect to translations, its projection on X is p, and the measure assigned to Ξ is one.

The translation invariance and the projection property follow immediately from the construction, specifically since p is invariant under the Markov transition map. To prove that $P_\infty(\Xi) = 1$, note that the complement of Ξ is the denumerable union of the subsets of X^∞ given by

$$\cdots \times X \times \cdots \times X \times D \times X \times \cdots \times X \times \cdots \tag{2.8}$$

where $D \subseteq X \times X$ occupies any two consecutive indices in the infinite string and is given by

$$D = \{(x, y) = y \notin F(x)\}. \tag{2.9}$$

Since $\operatorname{supp} \mu(x) \subseteq F(x)$ for p-almost every x, it follows that $P_2(D) = 0$. Hence the P_∞-measure of the set displayed in (2.8) is zero, and so is the denumerable union of these sets, namely the complement of Ξ. This completes the proof.

With the preceding five propositions, the proof of Theorem A is complete.

3. PROOF OF THEOREM B

PROPOSITION 3.1. *If the complete separable metric space X is locally compact and the set-valued function F has a closed graph, then an occupationally invariant probability measure p is subinvariant.*

Proof. Since X is separable and complete, the measure p is tight and hence for $C \subseteq X$ the value $p(C)$ is the supremum of $p(K)$ over the compact sets K satisfying $K \subseteq C$. This implies that in order to establish that p is subinvariant it is enough to verify (1.2) for compact sets. Furthermore, since X is locally compact, it follows that any compact set is the intersection of a sequence of compact sets whose boundary has p-measure zero. Since the inequality (1.2) is hereditary with respect to the intersection of a decreasing sequence, it follows that it is enough to establish (1.2) when K is compact and $p(\partial K) = 0$ (where ∂ denotes the boundary).

Let $(x_1, x_2, x_3, \dots)$ be a sequence which satisfies $x_{i+1} \in F(x_i)$, and whose occupational measures p_N (see (1.9)) converge to p as N converges to ∞. Since $x_{i+1} \in F(x_i)$ for each i, it follows that for any given K

$$P_N(K) \leq p_N(F^- K) + \frac{1}{N}. \tag{3.1}$$

If $p(\partial K) = 0$ then the left-hand side of (3.1) converges to $p(K)$; see, e.g., [5, p. 11]. Since K is compact and the set-valued map has a closed graph, it follows that $F^- K$ is closed. Hence (see [5, p. 11]) $\limsup p_N(F^- K) \leq p(F^- K)$. All in all, the limit of (3. 1) as N tends to ∞ verifies (1.2) for compact sets K with $p(\partial K) = 0$, and the proof is complete.

PROPOSITION 3.2. *If the complete separable metric space X is locally compact and the graph of F is closed, then a projectional invariant probability measure is an occupational invariant measure.*

Proof. Let p be a projectional invariant measure, and let P be the probability measure on X^∞, which is invariant under T, supported on Ξ, and whose projection is p. We consider the one-point compactification of

X, say $Y = X \cup \{\infty\}$, and extend F to Y by letting $F(\infty) = \{\infty\}$ and by adding ∞ to any value $F(x)$ when the latter is not compact. Then P is also a probability measure on the space of double infinite sequences Y^∞ of elements in Y, and it is invariant with respect to the translation operator T on Y^∞. Since T is continuous on Y^∞, the Choquet Theorem implies that P is in the convex hull of probability measures on Y^∞ which are extreme points in the space of probability measures invariant with respect to T; namely, a measure ρ exists on the space of invariant probability measures such that for every continuous function h: $Y^\infty \to R$ the equality

$$\int h(v) P(dv) = \int\left(\int h(v)\nu(dv)\right)\rho(dv) \tag{3.2}$$

holds, and the ρ-measure of the set of extreme points is 1. See, e.g., Pollicott and Yuri [10, Section 9.3]. It is easy to verify that each of the mentioned extreme points is ergodic; see [10, Lemma 9.4]. The compactness of the space and the continuity of T imply that each ergodic measure is occupationally invariant. The argument goes back to Kryloff and Bogoliuboff [8]. Indeed, let ν be ergodic. Then for each continuous real function h on Y^∞ the following convergence holds for ν-almost every v,

$$\lim \frac{1}{N+1}\sum_{i=0}^{N} h(T^i v) = \int h(v)\nu(dv). \tag{3.3}$$

The compactness implies that the space of continuous functions on Y^∞ is separable, hence for ν-almost every v the limit in (3.3) holds for every h, which means in particular that the occupational measures of the sequence of translations $T^i v$ converges weakly to ν. These arguments hold in Y^∞, but it is clear that if $\nu(X^\infty) = 1$ then ν is an occupational invariant measure in X^∞. Indeed, if (3.3) fails for a real function h which is continuous and bounded on X^∞, the tightness of P would imply that the convergence fails also for a modification of h which is continuous on Y^∞.

We claim that in (3.2) $\nu(X^\infty) = 1$ holds for ρ-almost every ν. Indeed, since P is tight on X^∞, if the claim fails it is easy to find a continuous function h on Y^∞ which gives large values to the complement of X^∞ and violates (3.2). Hence, the equality (3.2) holds in fact when the measure ρ is defined on the space of probability measures on X^∞, and it assigns measure 1 to the probability measures which are ergodic and invariant with respect to T. Since F has a closed graph, namely, Ξ is closed in X^∞, and since $P(\Xi) = 1$, we conclude that every measure in the support of ρ is supported on Ξ.

In view of the previous arguments, with the support of any ν participating in generating P (see (3.2)) an element $\chi \in \Xi$ can be found such that the occupational measures of the sequence $T^i\chi$ converge weakly to ν. Let

$$\chi = (\ldots x_{-1}, x_0, x_1, \ldots) \tag{3.4}$$

be this element. Then, the occupational measures of the sequence $(x_0, x_1, \ldots)$ converge weakly to the projection of ν. Since ν is invariant with respect to T we conclude that the projection p_0 of ν is an occupational invariant measure of F. Since taking convex combinations is preserved under the projection operator, the proof is complete.

With the preceding two propositions, the proof of Theorem B is concluded. Example 4.2 in the next section shows that the closed graph assumption cannot be dropped from the first proposition.

4. COMMENTS

As mentioned in the Introduction, a natural definition of a measure p being invariant, with respect to the set-valued map F is that p is invariant with respect to a measurable selection of F. The following example (borrowed from [1]) demonstrates why we have to settle for either the invariance with respect to a Markov transition map or selectable measures.

EXAMPLE 4.1. Let $X = [0,1]$ and let $F(x) = \{\frac{1}{2}x, \frac{1}{2} + \frac{1}{2}x\}$. The Lebesgue measure is an invariant measure of F. But the Lebesgue measure is not an invariant measure of any measurable selections of F. It is an invariant measure of the Markov transition map which, for a given x, assigns equal probabilities to the two elements in $F(x)$. A version of F which has a selection which makes the Lebesgue measure invariant is $F(x) = \{x, \frac{1}{2} + x\}$ when $0 \leq x \leq \frac{1}{2}$ and $F(x) = \{x - \frac{1}{2}, x\}$ otherwise.

What follows is the counterexample promised in the previous section.

EXAMPLE 4.2. Let $X = [0,1]$ and let $F(x) = [0,1]$ if $x > 0$ and $F(0) = \{1\}$. All the conditions required in the theorems are met except that F does not have a closed graph. Consider the sequence $(1, \frac{1}{2}, \ldots, \frac{1}{n}, \ldots)$. It clearly satisfies the conditions that establish the probability measure supported on $\{0\}$ as occupationally invariant. This measure is clearly not an invariant measure of F.

The various characterizations proposed in Properties I–V help to establish various properties of invariant measures of set-valued maps, including ergodic-type results. For instance, we can employ results in Vershik [13] which examine the ergodic properties of measures invariant with respect to a Markov transition map. The following would be a natural definition.

DEFINITION 4.3. A probability measure invariant with respect to the set-valued map F is ergodic if it is invariant and ergodic with respect to a Markov transition map m satisfying $\operatorname{supp} m(x) \subseteq F(x)$ for all x.

Since F is set-valued, the property that distinct ergodic (with respect to F) measures are orthogonal may not hold (this fact has a bearing on the applications; see [2]). The following is, however, true.

PROPOSITION 4.4. *A probability measure invariant with respect to the set-valued map F is decomposable into a direct integral of probability measures which are ergodic with respect to F.*

Proof. The proof follows directly from the definition and from Vershik [13, Proposition 4.2].

In the same spirit, the following results reproduce the existence result displayed in Aubin *et al.* [3] and, under conditions more relaxed than in [3], display a Poincaré recurrence theorem for set-valued maps.

PROPOSITION 4.5. *Let X be compact and let F be a set-valued map from X to subsets of X which has a closed graph. Then a probability measure subinvariant with respect to F exists.*

Proof. Since X is compact it follows that X^∞ is compact, and since F has a closed graph it follows that the subset Ξ of X^∞ is compact. Hence the translation operator T leaves Ξ invariant, and it is well known that a continuous map on a compact space has an invariant measure. Call it P. Its projection on X is a subinvariant measure.

PROPOSITION 4.6. *Let X be separable complete metric space and let F be a measurable set-valued map. Let p be a probability measure subinvariant, with respect to F. Then the following Poincaré recurrence property holds: for every Borel set B denote by B' the set of points x_0 in B for which a sequence $(x_0, x_1, \dots)$ exists such that $x_{i+1} \in F(x_i)$ and $x_j \in B$ for an infinite number of indices j. Then $p(B') = p(B)$. (In particular, it follows that if we denote $B_\infty = \bigcap_{n \geq 0} \bigcup_{m \geq n} F^{-m}(B)$ (here F^{-m} is the mth iteration of F^{-}), then $p(B) = p(B_\infty)$.)*

Proof. The Poincaré recurrence theorem holds for the operator T acting on (Ξ, P), where P is a probability measure on X^∞, invariant with respect to T, and whose projection is p and such that $P(\Xi) = 1$. Since Ξ consists of sequences satisfying the relation $x_{i+1} \in F(x_i)$, the recurrence stated in the proposition holds.

ACKNOWLEDGMENTS

I thank Anatoly Vershik for valuable discussions, and in particular for suggesting Property II as a definition of an invariant measure.

REFERENCES

1. Z. Artstein, Distributions of random sets and random selections, *Israel J. Math.* **46** (1983), 313–324.
2. Z. Artstein, Invariant measures of differential inclusions applied to singular perturbations, *J. Differential Equations* **152** (1999), 289–307.
3. J.-P. Aubin, H. Frankowska, and A. Lasota, Poincaré's recurrence theorem for set-valued dynamical systems, *Ann. Polon. Math.* **54** (1991), 85–91.
4. G. Beer, "Topologies on Closed and Convex Sets," Kluwer, Dordrecht, 1993.
5. P. Billingsley, "Convergence of Probability Measures," Wiley, New York, 1968.
6. C. Castaing and M. Valadier, "Convex Analysis and Measurable Multifunctions," Lecture Notes in Mathematics, Vol. 580, Springer-Verlag, Berlin, 1977.
7. F. Colonius and W. Kliemann, Some aspects of control systems as dynamical systems, *J. Dynam. Differential Equations* **5** (1993), 469–494.
8. N. Kryloff and N. Bogoliuboff, La théorie générale de la mesure dans son application à l'etude des systèmes dynamiques de la mécanique non linéaire, *Ann. Math.* **38** (1937), 65–113.
9. J. Neveu, "Mathematical Foundation of the Calculus of Probability," Holden–Day, San Francisco, 1965.
10. M. Pollicott and M. Yuri, "Dynamical Systems and Ergodic Theory," Cambridge Univ. Press, Cambridge, UK, 1998.
11. R. T. Rockafellar, Integral functionals, normal integrands and measurable selections, *in* "Nonlinear Operators and the Calculus of Variation," (L. Waelbroeck, Ed.), Lecture Notes in Mathematic, Vol. 543, pp. 157–208, Springer-Verlag, Berlin, 1976.
12. D. Ross, Selectionable distributions for a random set, *Math. Proc. Cambridge Philos. Soc.* **108** (1990), 405–408.
13. A. Vershik, Many-valued measure-preserving mappings (polymorphisms) and Markovian operators, *J. Soviet Math.* **23** (1983), 2243–2266.
14. B. Živaljević, Hyperfinite transversal theory, II, *Trans. Amer. Math. Soc.* **348** (1996), 1921–1938.